The Journal of Educational Issues of Language Minority Students
Volume 9, Fall, 1991 • Boise State University, Boise, ID 83725

Border College Consortium Descriptive Test of English Skills and Verbal Aptitude Test: Norming and Validation Study

Demetrius Karathanos

Introduction

Most of the U.S. communities located along the U.S.-Mexico border are heavily populated by persons of Mexican heritage. From Texas to California, it is not uncommon to find communities in which Mexican Americans comprise

---

Demetrius Karathanos, PhD, is a professor of management at Southeast Missouri State University in Cape Girardeau, Missouri.

as much as 90 percent of the population. The ethnic make up of these communities is also reflected in the student populations of the local public schools and institutions of higher education. Because of the proximity and easy access to Mexico, Mexican Americans along the border strongly identify with the Hispanic culture and language, and, in general, have limited opportunity to develop English skills. As a result, such students exhibit mild to severe deficiencies in the English language. These deficiencies are clearly reflected on test scores obtained from various standardized tests in English reading and writing (Karathanos, 1978). These scores tend to be grouped at the lower end of the scale and as such do not discriminate among students' abilities and potential for academic success. In addition, it is a well known statistical principle that when the spread of scores on a test is small, the reliability of the test is necessarily low (Nunnally, 1967). This problem is often referred to as the floor effect problem or as the truncated range problem. Thus, when the typical standardized English tests are administered to Mexican American students, the results are unreliable and of limited usefulness. Moreover, low scores of bilingual students on standardized tests have often been erroneously interpreted as indicating low potential for academic success with the result of denying equitable educational opportunity to many otherwise capable students.

These problems have been of great concern for many years to educators dealing with bilingual students (Karathanos, 1978). As an initial effort to provide solutions to such problems, the Border Community College Consortium undertook the project of developing two testing instruments to be used specifically with Mexican American students. The Consortium is made up of six community colleges along the U.S.-Mexico Border: Southwestern Community College and Imperial Valley College in California, Arizona Western College and Cochise College in Arizona, and Laredo Junior College and Texas Southmost College in Texas. Over a period of three years, a group of educators and students from the Consortium, in cooperation with other experts in bilingual education as well as testing specialists from the Educational Testing Service, developed two tests specifically for Mexican American students residing in the Southwest who plan to enroll in an institution of higher education: (a) The Descriptive Test of English Skills, and (b) The Verbal Aptitude Test. This project was financed in part by a grant from Education Professions Development Act, Title V-E, and a grant from the Title III, Division of College Support, U.S. Office of Education.

The following sections describe the two tests and present the results of the norming and validation study. Norming refers to the computation of various descriptive measures such as means, standard deviations, percentile ranks, and percentile bounds for the total tests and subsections of the tests. Validation refers

to the compilation and examination of various statistical characteristics which would help answer the question: "How well does this test measure what it is supposed to measure?" Such statistical characteristics usually include distribution characteristics, reliability coefficients, and correlation coefficients between the tests and other relevant variables.

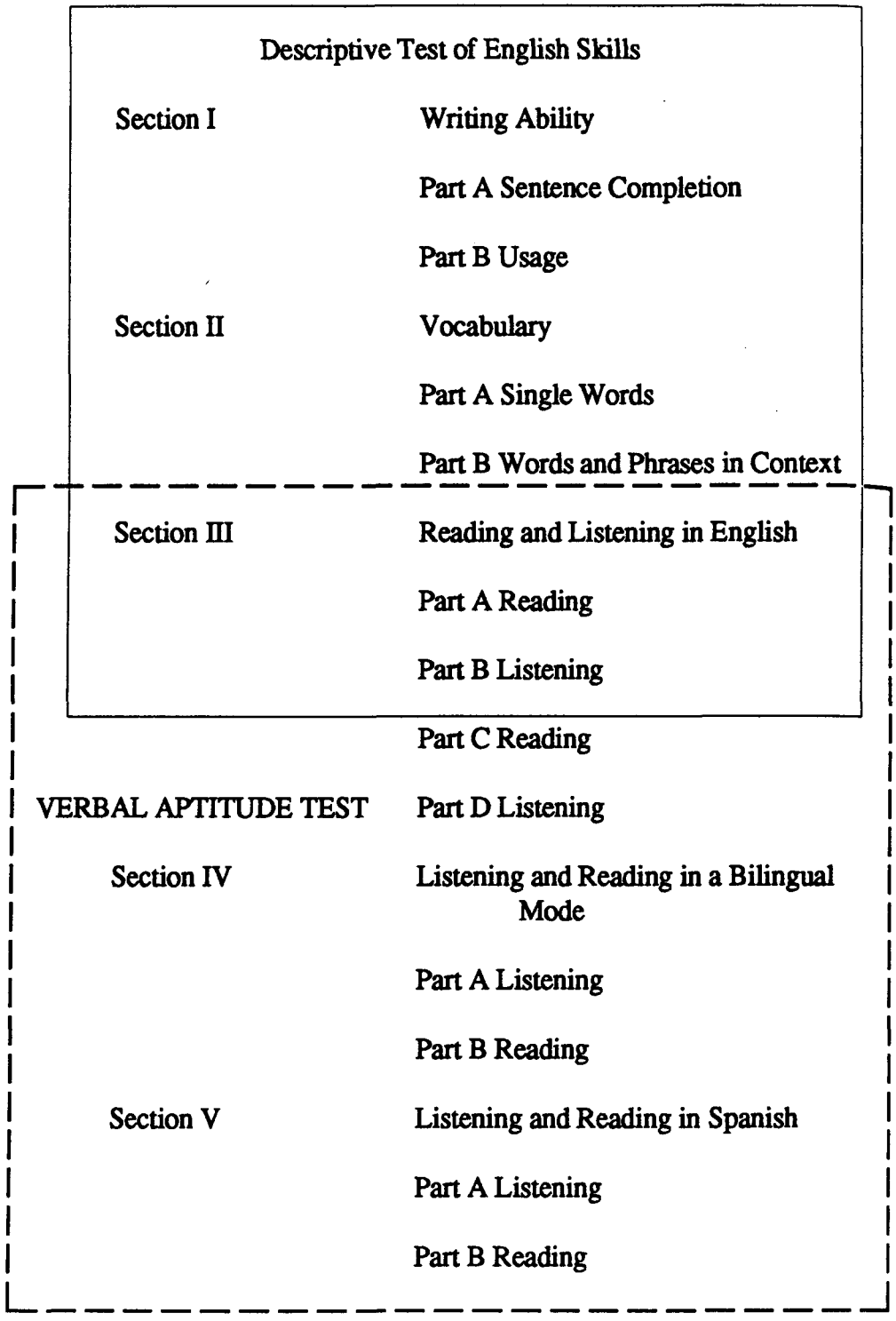
Description of the Tests

Figure 1 shows the structure of the Descriptive Test of English Skills and of the Verbal Aptitude Test. This diagram shows that parts A and B of Section III overlap both tests. When administered separately, the time required for the administration is one hour and 40 minutes for the Descriptive Test of English Skills and two hours and 50 minutes for the Verbal Aptitude Test. However, when both tests are administered at one time, the overlapping of parts A and B of Section III results in reducing the total administration time to three hours and 30 minutes. The tests are entirely group-administered. Except for the listening portions (Section III-Parts A and D, Section IV—Part A, and Section V—Part A), the content of the tests, including the instructions, is read by the students. The listening portions, including the instructions, are presented to students via a tape player.

The Descriptive Test of English Skills was designed to be used primarily in two ways: (a) as a diagnostic tool through which specific deficiencies in the English language could be identified, and (b) as a placement instrument through which students can be placed in English classes of the appropriate level. The Verbal Aptitude Test is composed of a section in English, a unique section in a bilingual mode (a mixture of English and Spanish commonly used along the border), and a section in Spanish. This test was designed to do the following: (a) provide guidance and placement information to students of Spanish speaking backgrounds and to their counselors, and (b) provide an indication of potential for future academic success. In planning a student's program, it is important to have an indication of that student's verbal skills in comparison with those of other students and to know in which language the student's present strength lies.

The "Guide for Teachers and Counselors" and the "Guide for Students" contain detailed information regarding the structure of the two tests and are available from Imperial Valley College, Imperial, CA.

Figure 1. Diagram of content of the two tests



Norms

Population

The population on which norms were derived for the tests was defined as all first semester Mexican American freshmen in the consortium colleges registered for seven or more units.

Sample

During the period of the third through the fifth week of the fall semester, the two tests were administered to a sample of 804 first semester Mexican American (self-identified) students in five consortium colleges (174 at Southwestern College, 225 at Imperial Valley College, 48 at Arizona Western College, 267 at Laredo College, and 90 at Texas Southmost College). These students were selected as follows: In those schools having populations numbering fewer than 500, all students in the population were invited to take the tests. In schools with populations over 500 students, 500 students were randomly selected and were invited to take the tests. The proportions of respondents ranged from a high of 65 percent in one school to a low of 20 percent in another. The overall rate of response was 40 percent. The wide range of responses can be primarily attributed to the fact that, although the consortium colleges share many common characteristics, there are wide differences in institutional philosophies. For example, admissions policies vary significantly among schools, ranging from an open door policy to a policy of selective admissions. The school with the lowest response rate has a no-testing policy whereby a student is not required to take any tests whatsoever. Obviously, a weakness in this study is the lack of a truly random sample. However, it should be recognized that under the present operating conditions of the consortium colleges, it would be impossible to administer the tests to a truly random sample of adequate size. For the purpose of determining what effects, if any, the lack of a random sample may have had on the obtained results, the respondents were compared to the nonrespondents in one school (Imperial Valley College) on selected variables at the conclusion of the fall semester. The results are shown in Table 1.

It appears that the two groups exhibit differences on several variables. The students who took the test seem to have a better achievement record as shown by the higher overall GPA. They also appear to have a better persistence record as shown by several factors: (a) the higher percentage of those completing seven

or more units, (b) by the higher average number of units completed, and (c) by the lower percentage of complete withdrawals.

Table 1. Comparison of Students Who Took the Tests and Those Who Did Not at Imperial Valley College

	1st Semester Overall GPA	1st Semester GPA in English	Percent Completing 7 or more Units	Average number of Units Completed	Percent Who Withdrew Completely
Students who took the test	x = 2.315 S = .714 N = 194	x = 2.448 S = .924 N = 112	80%	x = 12.30 S = 4.50 N = 201	9%
Students who did not take the test	x = 2.174 S = .846 N = 84	x = 2.270 S = .882 N = 34	49%	x = 10.30 S = 5.19 N = 94	27%
Two-tail t-test p<	.10	.15		.001	
Two-tail binomial test p<			.001		.001

Although the results of Table 1 apply only to students at Imperial Valley College, I suspect that similar differences between the two groups might exist in the other schools as well. When interpreting scores on these tests, the above results should be taken into consideration.

Score Interpretation

To avoid the obvious difficulties of interpreting raw scores, percentile ranks and percentile bands were computed for interpreting performance on these tests. Percentile rank and band tables for nine sections of the two tests are available

at Imperial Valley College, Imperial, CA.

Precision of Norms

The precision of the reported percentile ranks for the two tests depends primarily on the size of the standard error of measurement of the test. This error of measurement is a direct function of the reliability coefficient and provides an indication of how representative the items used on the test are of the total possible items in the area covered by the test (Nunnally, 1967). Since the standard error of measurement is used in constructing the percentile bands, these bands account for the error due to the sampling of items.

One additional source of possible error not incorporated in the percentile bands is the lack of a truly random sample as discussed earlier in this section. This should be kept in mind whenever interpreting scores on these tests.

Statistical Characteristics

Score Distribution

Figure 2 shows the distribution of the scores of the entire Descriptive Test of English Skills, and Figure 3 shows the distribution of scores on Section III — Parts B and D (total listening score in English). Both distributions show good dispersion with no floor effect. The distributions of all other scales are similar to those shown in Figures 2 and 3.

Table 2 shows means and standard deviations for all major sections and subsections of the two tests.

Table 3 shows the relative performance on 20 scales for men and women. Only two scales show significant differences.

Reliability

Table 2 shows the reliability coefficients for 20 scales and subscales which have 15 or more items. These coefficients are internal consistency measures and were computed by the Kuder-Richardson formula 20 (Nunnally, 1967). Reliability coefficients for scales with fewer than 15 items are not reported for two reasons: (a) As a rule of thumb, fewer than 15 items was considered too small, and (b) these scales correspond to very specific, narrow skills.

Figure 2. Score Distribution on the Total Descriptive Test of English Skills, N = 804

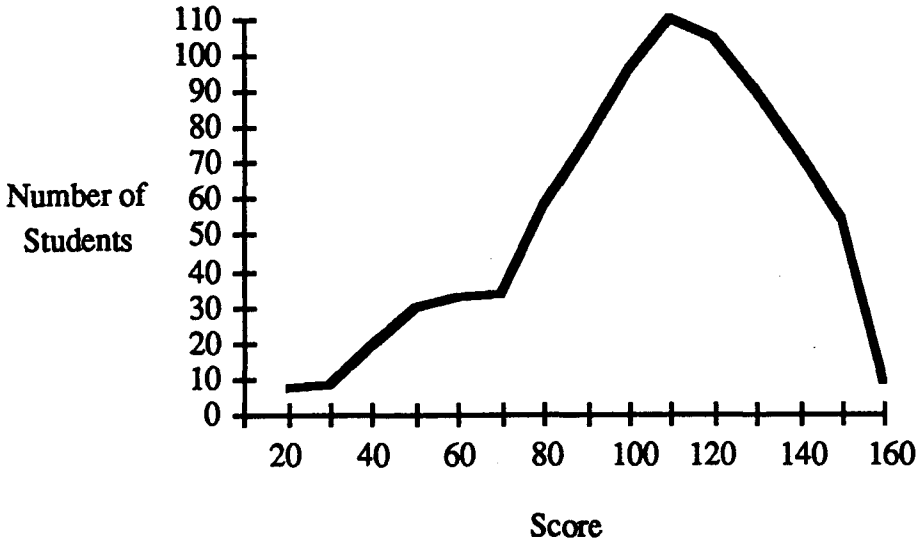


Figure 3. Score Distribution on Section III — Parts B and D (Total Listening Score in English), N = 804

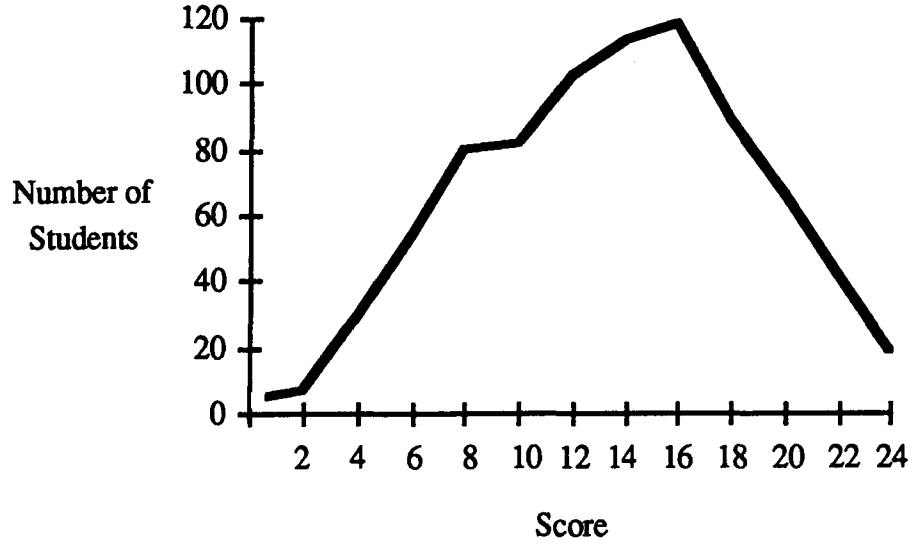


Table 2. Means, Standard Deviations, and Reliability coefficients for Various Sections and Subsections of the Tests (Total Consortium, N = 804)

Description of Section or Subsection	Number of Items	Mean	Std. Dev.	Rel. Coeff.
1. Descriptive Test of English Skills	175	115.1	31.1	.97
2. Section I - Writing Ability	65	46.9	13.1	.95
3. Part A (Section I) Usage	30	24.1	5.4	.90
4. Part B (Section I) Usage	35	22.8	8.1	.92
5. Verbs	13	9.6	3.2	NA
6. Pronouns	9	6.9	2.1	NA
7. Prepositions	10	7.1	2.0	NA
8. Adverbs	6	4.2	1.3	NA
9. Nouns	4	3.1	1.0	NA
10. Noun Predictors	5	4.3	1.0	NA
11. Minimal Pairs	2	0.9	0.8	NA
12. Plurals	3	1.9	1.1	NA
13. Comparatives	3	2.1	1.0	NA
14. Apostrophes	3	1.8	1.0	NA
15. Idioms	3	2.1	0.9	NA
16. Tag Questions	2	1.1	0.8	NA
17 Double Negatives	2	1.8	0.5	NA
18. Section II - Vocabulary	60	38.3	10.9	.92
19. Part A (Section II) - Single Words and Phrases	40	23.8	7.6	.88
20. Part B (Section II) Words and Phrases in Context	20	14.5	3.9	.81
21. Section III - Total Verbal Aptitude in English	80	47.9	14.4	.93
22. Parts A & C (Section III) Total Reading Score	40	23.6	7.6	.87
23. Parts B & D (Section III)	40	24.2	7.7	.88

Total Listening Score				
24. Part A (Section III) Partial Reading Score	25	14.1	5.1	.82
25. Main Idea	3	1.6	1.0	NA
26. Explicit Detail	9	5.8	2.0	NA
27. Implication	13	6.7	2.9	NA
28. Part B (Section III) Partial Listening Score	25	15.9	5.0	.82
29. Main Idea	2	1.1	0.7	NA
30. Explicit Detail	12	7.3	2.4	NA
31. Implication	11	7.5	2.5	NA
32. Part C (Section III) Partial Reading Score	15	9.6	3.1	.74
33. Part D (Section III) Partial Listening Score	15	8.4	3.3	.77
34. Section IV - Total Bilingual Mode Score	40	23.6	7.0	.85
35. Part A (Section IV) Bilingual Listening	25	14.6	4.0	.73
36. Part B (Section IV) Bilingual Reading	15	8.9	3.8	.83
37. Section V Total Spanish Score	60	33.8	12.1	.93
38. Part A (Section V) Spanish Listening	30	16.7	5.8	.83
39. Part B (Section V) Spanish Reading	30	16.7	5.8	.83

Table 3. Means and Standard Deviations on Selected Sections and Subsections for Males (N=349) and Females (N=455) (Total Consortium)

Description of Section or Subsection	Mean		Std. Deviation	
	Male	Female	Male	Female
1. Descriptive Test of English Skills	114.3	115.8	31.5	30.7
2. Section I - Writing Ability	45.9	47.5	13.5	13.0*
3. Part A (Section I) Sentence Completion	23.7	24.3	5.9	5.5
4. Part B (Section I) Usage	22.3	23.3	8.2	8.0*
5. Section II Vocabulary	38.0	38.5	11.0	10.9
6. Part A (Section II) Single Words and Phrases	23.7	24.0	7.6	7.6
7. Part B (Section II) Words and Phrases in Context	14.3	14.5	4.0	3.9
8. Section III Total Verbal Aptitude in English	47.8	47.9	14.6	14.5
9. Parts A & C (Section III Total Reading Score	23.5	23.7	7.7	7.5
10. Parts B & D (Section III) Total Listening Score	24.3	24.2	7.6	7.8
11. Part A (Section III) Partial Reading Score	14.0	14.1	5.3	5.1
12. Part B (Section III) Partial Listening Score	15.7	15.4	5.3	5.3
13. Part C (Section III)	9.5	9.6	3.2	3.0
14. Part D (Section III)	8.3	8.5	3.3	3.4
15. Section IV Total Bilingual Mode Score	23.3	23.7	7.1	6.9
16. Part A (Section IV) Bilingual Listening	14.4	14.8	4.0	4.0
17. Part B (Section IV)	8.9	8.9	3.9	3.8

Bilingual Reading				
18. Section V Total Spanish Score	33.1	34.4	12.8	11.6
19. Part A (Section V) Spanish Listening	16.4	16.9	6.0	5.7
20. Part B (Section V) Spanish Reading	16.7	17.5	7.4	6.6

*significant (2-tail t-test, $p<.10$)

Validity

A. Descriptive Test of English Skills

The Descriptive Test of English Skills (DTES) was primarily designed to be used as a diagnostic instrument and as a placement tool. No formal studies have been conducted to assess the diagnostic utility of the DTES. However, informal observations of teachers of English as a second language who have used the DTES indicate that the test is a useful diagnostic tool. More specifically, they have found Section I—Part B (Usage) and Section II — Vocabulary, particularly helpful.

Lopez (1981) investigated the placement efficiency of the Descriptive Test of English Skills. Using a sample of 376 first semester Mexican American students at Imperial Valley College, he used three predictor variables—Descriptive Test of English Skills (DTES), Nelson Denny Reading Test (NDRT), and High School Grade Point Average (HSGPA), to predict grades in five levels of English writing courses. His results showed that the DTES was the best predictor for the lower level courses while the HSGPA was the best for the higher level courses. The NDRT was generally found to be a poor predictor at all levels. Because of differences in course content in various English programs, placement studies cannot be generalized but must be conducted at each individual school using the test. Several additional indications are encouraging regarding the validity and the placement utility of the DTES:

1. The wide score dispersion and the fairly high standard deviations in all scales indicate that the tests provide good discrimination among the target students.

2. The correlations shown on Table 4 between the Border Consortium tests

and other standardized English tests at Southwestern Community College and Laredo Junior College range from .54 to .73. These moderate correlations are of a desirable magnitude since they are high enough to indicate that all of the tests measure similar skills (English), but they are not high enough to make the tests interchangeable.

3. Table 5 shows correlations between grades in specific English courses and scores on various portions of the DTES at Laredo College. English 01 is a general skills course designed to bring a student's skills to the level required for the standard English composition course. English 21 is the typical English composition course. These correlations of low to moderate magnitude should provide a basis for further research on the placement efficiency of the DTES and on the establishment of concurrent validity.

B. Verbal Aptitude Test

Table 6 shows intercorrelations between the three distinct parts of the Verbal Aptitude Test—English, Bilingual Mode, and Spanish, computed on the total sample. The low correlation of .23 between English and Spanish indicates that high aptitude in one language associates weakly, but significantly (p .01), with high aptitude in the other. The moderate and statistically significant (p .01), correlations of the Bilingual Mode with English and Spanish (.64 and .62, respectively) seem reasonable since the Bilingual Mode is a mixture of the two languages. Table 7 provides correlations between various parts of the Verbal Aptitude Test and three performance variables: (a) first semester overall GPA, (b) first semester GPA in English, and (3) grade in Spanish 41 (an elementary Spanish course).

The correlations in the first two columns of Table 7 range between .15 and .31 and show weak but significant (p .01) association between the various parts of the Verbal Aptitude Test and the criteria of first semester overall GPA and GPA in English. These correlations are not particularly strong evidence of predictive validity. Further research, perhaps using GPAs based on longer periods, is desirable. The correlations in the third column range between .33 and .57 and show stronger, as well as statistically significant (p .01), association between Bilingual and Spanish Aptitude and the criterion of performance in an elementary Spanish course. These correlations provide somewhat stronger evidence of predictive validity, but it should be emphasized that the criterion variable (performance in an elementary Spanish course) is rather narrow in scope.

Table 4. Correlations Between Border College Tests and Other Standardized English Tests (All Correlations Significant at the .01 Level)

English Proficiency Tests	N	Border Consortium Tests					
		DTES			Verbal Aptitude		
		Total	Writing	Vocab	Reading	Listening	Total
Nelson Denney Vocabulary*	143			.67			
Nelson Denney Comprehension*	143				.56	.46	
Nelson Denney Total*	143	.67	.54				
College English Placement Test*	143	.73	.64	.72	.69	.68	
ACT English**	200	.70	.65	.67	.65	.59	.65

*Southwestern Community College
**Laredo College

Table 5. Correlations of Grades in English 01 and English 21 with Scores on Portions of the DTES, (Laredo College) (All Correlations Significant at the .01 level)

	N	DTES 1*	WRIT 2	SENT 3	USAGE 4	VOCAB 5	WF-S 6	WF-C 7	READ 8	LIST 9
English 01	130	.44	.41	.39	.38	.42	.42	.33	.41	.33
English 21	93	.47	.45	.44	.40	.46	.41	.44	.39	.30

Table 5, cont.
* 1. Descriptive Test of English Skills; 2. Writing Ability (Section I); 3. Sentence Completion (Part A-Section I); 4. Usage (part B-Section I); 5. Vocabulary (Section II); 6. Single Words and Phrases (Part A-Section II); 7. Words and Phrases in Context (Part B-Section II); 8. Reading (Parts A & C-Section III); 9. Listening (Parts B & D-Section III)

Table 6. Intercorrelations Among the Three Parts of the Verbal Aptitude Test, N=804 (All Correlations Significant at the .01 Level)

	English	Bilingual	Spanish
English			
Bilingual	.64		
Spanish	.23	.62	

Table 7. Correlations of (1) First Semester Overall GPA, (2) First Semester English GPA, (3) Grade in Spanish 41 With Various Parts of the Verbal Aptitude Test (All Correlations Significant at the .01 Level)

	First Semester College GPA N=651	First Semester English GPA N=366	Grade In Spanish 41 N=81
English Aptitude	.28	.30	
English Reading	.30	.31	
English Listening	.22	.24	
Bilingual Aptitude	.26	.26	.44
Bilingual Reading	.29	.31	.46
Bilingual Listening	.18	.16	.33
Spanish Aptitude	.18		.55
Spanish Reading	.15		.48
Spanish Listening	.15		.57

Summary

This paper describes the background and the rationale for the development of two new tests and presents the results of the norming and validation study on these tests.

The Descriptive Test of English Skills (DTES) and the Verbal Aptitude Test (VAT) were developed by the Border Community College Consortium with assistance from the Educational Testing Service. These tests were designed specifically for bilingual Mexican-American students with the objective of overcoming the inherent inadequacies of the typical standardized college entrance English tests when administered to this group of students.

Norms in the form of percentile ranks and percentile bands have been computed for the two tests and for seven subsections of the test. These norms are based on scores obtained from a sample of 804 students.

Frequency distributions for two scales have been presented showing good score dispersion. The distribution of all other scales is similar to the two presented. The dispersion of these scales shows that the floor effect problem is absent in these tests in contrast to other standardized English tests.

Means, standard deviations, and internal consistency reliability coefficients for the overall tests and their major parts have been presented. These coefficients range from 0.73 to 0.97. In general, the higher coefficients correspond to the longer scales as would be expected.

Comparisons of males and females on mean scores in 20 scales show some small differences which do not give rise to any major concern.

Correlations between various parts of the Consortium tests and parts of other standardized English tests have been presented as evidence of concurrent validity for the two tests.

Correlations between performance in two levels of English courses and scores on various parts of the DTES have been presented as evidence of placement efficiency of the DTES. In addition, results obtained by another investigator on the placement efficiency of the DTES have been reported.

Finally, correlations between three measures of academic success and various scales of the VAT have been presented as evidence of predictive validity for this test.

As noted earlier, because bilingual Mexican American students tend to cluster at the lower end of the range of scores on typical standardized English tests, the reliability of these tests is inherently low. Since high reliability is a necessary condition for good validity, these tests are generally not valid when administered

to bilingual Mexican American students. The results presented in this study collectively provide evidence that the DTES and the VAT are reliable and valid for these students. As such, these tests appear to have overcome the inadequacies of other standardized English tests and seem to be viable alternatives for bilingual Mexican American students. Details on how these tests can be used by appropriate college personnel can be found in the "Guide for Teachers and Counselors" which is available at Imperial Valley College, Imperial, CA.

It should be noted, however, that additional research related to the diagnostic and placement efficiencies of the DTES and to the predictive validity of the VAT is needed.

References

Karathanos, D. (1978). *Border college consortium testing program for bilingual students.* Paper presented at the California Community and Junior College Association Research and Development Conference, Pacific Grove, CA, April 1978.

Lopez, Hector, L. (1981). The effectiveness of the descriptive test of English skills and other selected variables in predicting academic performance in English courses for Mexican American students (Doctoral dissertation, Northern Arizona University, 1981). *Dissertation Abstracts International,* 42, 149A.

Nunnally, J. C. (1967). *Psychometric Theory.* New York: McGraw-Hill.